A REVIEW OF THE CULTURAL HERITAGE LINKED OPEN DATA ONTOLOGIES AND MODELS

F. Liu, J. Hindmarch, M. Hess

University of Bamberg, Digital Technologies in Heritage Conservation, Institute of Archaeology, Heritage Conservation Studies and Art History, Germany – (fangchao.liu, john.hindmarch, mona.hess)@uni-bamberg.de

KEY WORDS: Ontologies, CIDOC CRM, Europeana, Sampo Model, Linked Open Data.

ABSTRACT:

In recent years, there has been a trend towards synergies among cultural heritage institutions, such as GLAM (Galleries, Libraries, Archives & Museums) collaborations in the field of information resource management. These institutions have been actively promoting the construction of digital humanities projects. Such projects primarily target data resource construction and digital humanities services, which are in line with the service functions of GLAM institutions. With the development of networking and digitisation, museums have been making more prominent contributions to knowledge sharing and dissemination.
Digital technology has brought about profound changes in the academic research environment. The expansion of traditional research and the emergence of new questions in the digital environment have promoted novel perspectives, thinking, and methods. With the increase in digitisation of cultural heritage, the scale of data has grown rapidly, and the types of data available have been continuously enriched, from knowledge aggregation to knowledge organization and even knowledge transmission. Therefore, it is necessary to examine the tools of digital humanities, which constitute an important means of knowledge production, and to work with traditional GLAM institutions to promote digital thinking and interdisciplinary knowledge integration methods. At the technical level, methods such as knowledge graphs and ontologies are used to construct a multi-layered and systematic knowledge system. This paper presents a review focused on three cultural heritage models, CIDOC CRM, Europeana and the Sampo Model, and reflects on three aspects of the current models: technical framework construction, digital resources, and service systems.

1. INTRODUCTION

1.1 General Introduction

Cultural and GLAM institutions use many different metadata schemes, and thus to efficiently organise and manage metadata of different standards, uniform data rules need to be established. Although digitising cultural heritage can increase access and usability for people with widely differing needs, as the volume of data increases, so does the need for an effective data model. This review focuses on three cultural heritage models, CIDOC CRM, Europeana and the Sampo Model. We compare the current situation of the models in three aspects: technical framework construction, digital resources, and service systems. All the models selected in this review focus on long-term storage and providing persistent access to large digital resources.

1.2 CIDOC CRM

The CIDOC Conceptual Reference Model (CIDOC CRM) is an ISO standard ontology (ISO 21127:2014) used to enable the exchange of information and meaningful interoperability between GLAM and other cultural institutions (Čerāns et al., 2021). It was developed by the International Committee for Documentation (CIDOC) under the International Council of Museums (ICOM). It assists researchers, managers and the public in related fields in cultural heritage information by providing a common and extensible semantic framework (Fig 1). It is a formal ontology that provides a framework for describing cultural heritage objects and their relationships to each other in a structured and consistent way. It is designed to be a conceptual tool that can be used to represent and manage the complex relationships between different types of objects in cultural heritage collections.
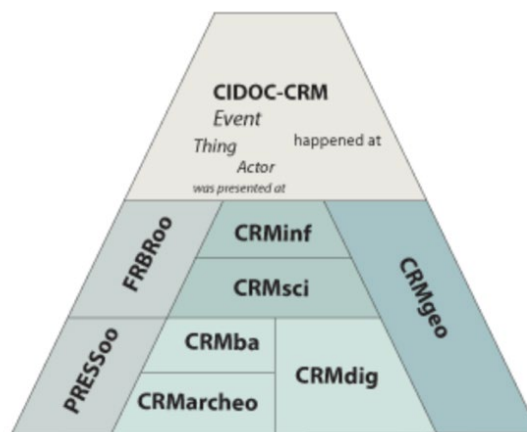

Figure 1. CIDOC CRM family of Models

1.3 Europeana

In order to enable European stakeholders to better understand their own historical treasures and cultural heritage, in 2005, 19 countries, led by France, began creating a digital library, which first opened in 2008. Originally the "European Digital Library", as more and more museums, galleries and other institutions joined, it was renamed Europeana (Latin, meaning "things European") (Purday, 2009a). Currently, Europeana includes data resources from around 4,000 cultural institutions. Europeana designed the Europeana Semantic Element metadata scheme (ESE) based on its own data characteristics. With the development of Linked Data, Europeana created the open Europeana Data Model (EDM) with reference to relevant standards and specifications such as METS (Metadata Encoding and Transmission Standard) and RDF (Resource Description

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

Framework). In terms of metadata semantic enrichment, the EDM provides a data model basis for semantic association and description between different institutions. Europeana focuses on long-term storage and persistent access to massive digital resources and is committed to the global sharing of digital resources within the scope of copyright licenses.

1.4 Sampo Model

Sampo model is a meta-model for creating cross-domain ontology and linked data integration. The Sampo model creates semantic portals using principles of re-usability (Hyvönen, 2022) and can link and make interoperable different types of resources (Fig 2). The Sampo model includes three main parts: a model for creating and publishing linked data; a resource content view for end users; and a faceted retrieval and data analysis based on web portals. In addition to providing services such as text retrieval, browsing and downloading of traditional resources, it also provides a variety of content views to help users perform multi-dimensional semantic retrieval and analysis. Based on the Sampo model, several thematic semantic portals have been developed, such as WarSampo, CultureSampo, AcademySampo, FindSampo, etc.

Figure 2. Sampo Model based on a shared ontology (Hyvönen, 2012)

1.5 Summary of the introduction

With the massive increase in the availability of digital resources, integration and exploration options have become a public need. In order to address this demand, countries have developed and established digital resource discovery platforms to realise the shared application and value-added development of digital resources. To facilitate reuse and ontology layering, the ontology engineering community has suggested various levels of abstraction, differentiating four distinct degrees of abstraction in ontology design, as indicated in Fig 3 (Haller & Polleres, 2020).
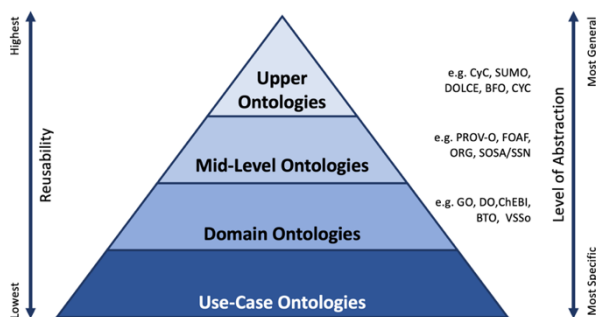

Figure 3. Levels of Abstraction in Ontology Design

With the help of the network of digital systems, the digitised data stored in heritage institutions can be shared not only within the specialist community itself, but also enables the involvement of the public (Arera-Rütenik, 2021). While serving the public, all three models in this review are committed to providing access to digital resources only within the scope of copyright licenses. Different platforms can gather abundant digital resources from cultural heritage institutions and provide bundled access, comprehensively enhancing the level of digital services provided to the public.

2. CIDOC CRM

CIDOC CRM stands for "Conceptual Reference Model" developed by a committee of ICOM, CIDOC (see section 1.2). It is based on the idea of an "object-oriented" model, where each object in the collection is treated as an individual entity, with its own unique set of properties and relationships to other objects. Widely used in the field of cultural heritage documentation and management, CIDOC CRM provides a standard vocabulary and framework that can be used to describe and exchange information about cultural heritage objects, regardless of the specific system or software used to manage the collection.

To quantify the academic publishing rate on CIDOC CRM as a tool we use CiteSpace to visualise the knowledge graph of publications (Chen, 2004 & 2006). Following are some basic analyses of publications on CIDOC CRM. The database used is the Web of Science Core Collection, and the search topic TS=(CIDOC CRM).

Figure 4. Publications on CIDOC CRM (Created at 20.03.2023)

The line chart shows the annual publications, and the bar chart the total publications. It can be seen that the first article was published in 1999, with the first peak in 2010 with 11 papers, and the highest number of publications in 2015, with 18 articles. The past several years (excluding 2020) have been relatively stable, with around 15 articles per annum. This indicates that the CIDOC CRM model is well established and still an object of active research. The Keywords Cluster analysis of CIDOC CRM from Citespace is shown in Fig 5 and the Timeline in Fig 6. The analysis results have six clusters; CIDOC CRM, semantic web, digital humanities, historic building, knowledge base & data standards. As can be seen in Fig 6, digital humanities and data standards appeared first, but data standards lasted only a short time and did not continue after 2017. Similarly, knowledge base also disappeared around 2019. Historic building appeared latest in 2018, while the other topics have been continuous throughout the period. This suggests that "data standards" and "knowledge base" are no longer the focus of research, or perhaps have been

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

merged into other clusters. On the other hand, CIDOC-based "historic buildings" have been a research hotspot in recent years.
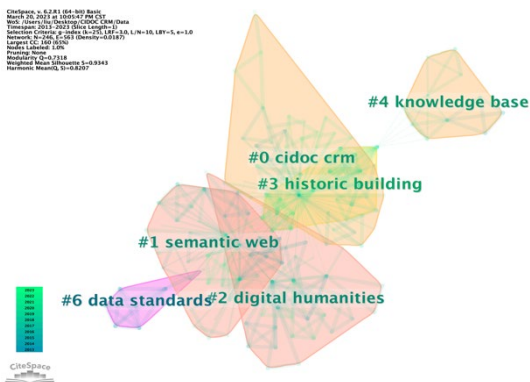

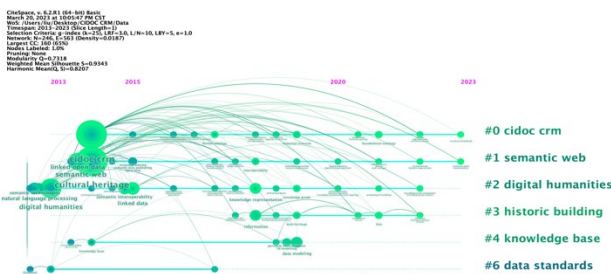**Figure 5.** Cluster Map (Created at 20.03.2023)

Figure 6. Timeline of Cluster (Created at 20.03.2023)

2.1 Technical Framework Construction

CIDOC CRM is divided into two main parts: the "core" model, which defines the basic concepts and relationships used to describe cultural heritage objects, and a set of "extensions" that provide additional concepts and relationships specific to different types of cultural heritage collections, such as archaeological, ethnographic, or museum collections.

The main function of the CIDOC CRM is to provide a framework for the mediation of cultural heritage information and, in doing so, to act as the semantic "glue" that unites various localised sources of information into a valuable and cohesive global resource (Tzitzikas et al., 2022). CIDOC CRM is based on a formal ontology and the technical framework of CIDOC CRM includes the following elements: Classes, Properties and Relationships, Formal Language and Extension.

CIDOC CRM defines a set of classes that are used to represent different types of entities in the cultural heritage domain. These classes provide a formal structure for organizing and describing cultural heritage data, such as "E21 Person" for people, and "E71 Human-Made Thing" for objects that have been created by humans (Mazurek et al., 2012). Subclasses further refine the classification. These classes provide a formal structure for organising and describing cultural heritage data. Fig 7 shows the CIDOC CRM Top level classes. Alongside Classes, Properties are used to represent integrated information in a rich semantic network for CIDOC CRM compliant datasets (Tzitzikas et al., 2022). Properties describe the various characteristics of cultural heritage objects, for example "P1 is identified by", "P2 has type", and "P46 is composed of". Version 7.1.2 of the CRM includes 81 Classes and 160 Properties (*Classes & Properties Declarations of CIDOC-CRM Version: 7.1.2*, 2023).
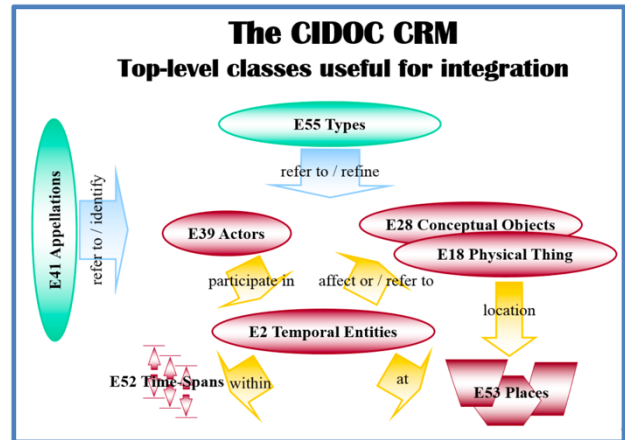

Figure 7. The CRM Top Level (Oldman, 2014)

Relationships are used to describe the connections between cultural heritage Classes and Properties. It is provided by CIDOC CRM as a formal language and structure for representing cultural heritage data, including objects, events, actors, and places. The current formal language of CIDOC CRM is Web Ontology Language (OWL), a standard for representing ontologies on the web. CIDOC CRM is also related to other ontologies, like Dublin Core, FOAF, and SKOS. These define concepts and relationships that are complementary to CIDOC CRM and can be used to extend its capabilities.

The community has contributed "extensions" to the ontology for, for example, the description of conservation activities, paintings and other specific cases. It can be extended programmatically, creating new Classes and Properties on top of the core model to provide a domain that is more suited to a particular knowledge organisation (Niccolucci, 2017). For example, the extension "CRMgeo" is focused on the spatiotemporal properties of temporal entities and persistent items, enabling information integration on a spatiotemporal level based on the semantics defined in CIDOC CRM (Hiebel et al., 2017).

2.2 Digital Resources

The digital resources of CIDOC CRM include a variety of materials that support the use and implementation of the conceptual reference model. These resources are designed to help users understand and apply the model in different contexts and to promote interoperability and standardisation in the management of cultural heritage collections.

The CIDOC CRM ontology formally represents the model in a machine-readable format. The ontology defines the classes, properties and relationships of the model, and allows for the exchange of information between different systems and applications. The ontology is available in RDF, OWL, and XML etc., and is used by software applications and services to represent, query, and reason about cultural heritage data. A variety of applications for CIDOC CRM have been developed (*Use Cases | CIDOC CRM*, 2023). There is also a user guide, a technical specification, and a set of guidelines for implementing the model. The documents provide detailed information about the concepts, properties, and relationships defined in the model and instructions for using the model in different applications.

2.3 Service System

The service system of CIDOC CRM refers to the set of services, applications, and tools built on top of the conceptual reference

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

model to provide specific functionality and support for cultural heritage management. These services and applications are designed to work with CIDOC CRM as a foundation and to provide value-added functionality for different use cases and contexts.

Many collection management systems for GLAM institutions use CIDOC CRM as the underlying data model for managing and describing cultural heritage objects. Tools for cataloguing, indexing, retrieval, and analysis allow users to work with data compliant with the CIDOC CRM ontology. It has also been adopted as an exchange standard for sharing cultural heritage data across different systems and organizations. These standards include the CIDOC CRM XML schema, which provides a common format for exchanging CIDOC CRM data, and the Linked Open Data (LOD) standard, which allows CIDOC CRM data to be integrated with other cultural heritage data sources on the web. These applications use the CIDOC CRM ontology as a foundation for describing cultural heritage objects and allow users to discover and explore relationships between different objects and resources. The service system of CIDOC CRM enables users to apply the model in a wide range of contexts and use cases by providing a common foundation for different services and applications.

CIDOC CRM involves a range of activities and processes aimed at developing, maintaining, and promoting the conceptual reference model. These activities are coordinated by ICOM, which oversees the development of the model and provides guidance and support to the community of users and developers.

3. EUROPEANA

Europeana is a digital platform that aggregates and makes accessible cultural heritage content from across Europe. It was launched as an initiative of the European Commission with the goal of providing a single access point to the wealth of digital cultural heritage resources held by European libraries, archives, museums, and other cultural institutions. Europeana aggregates information from digital collections and thus provides access to digital objects including books, manuscripts, photographs, paintings, films, and audio recordings, which, in the past, was difficult or in some cases impossible (Purday, 2009b). The platform is designed to make cultural heritage more accessible, usable, and reusable for researchers, educators, and the general public. Europeana also supports the development of digital skills and competencies in the cultural heritage sector and works to promote the use of digital technologies in cultural heritage preservation and dissemination.

Following are some basic analyses of Europeana from Citespace. The database is the Web of Science Core Collection, and the search topic TS=(Europeana).

In Fig 8, the line shows the annual publications, while the bars show total publications. The first article on the topic 'Europeana' was published in 2009, and there followed an upward trend, peaking in 2014 with 25 articles. Subsequently the volume was relatively stable, before showing a sharp decline from 2021 onwards. The period from 2013 to 2020 was the most popular period for studying Europeana.
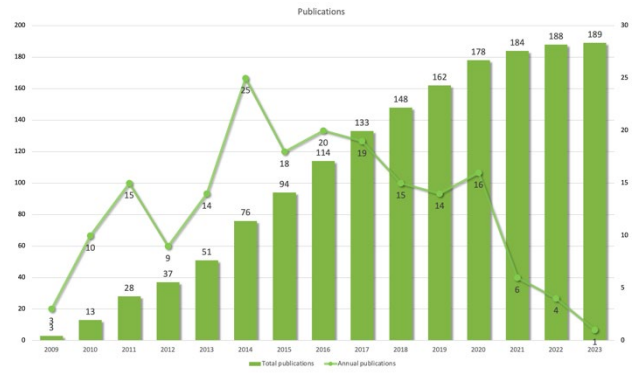

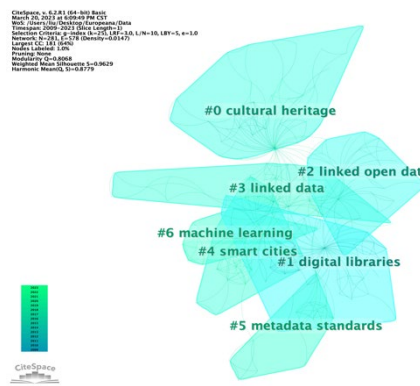**Figure 8.** Publications of Europeana (Created at 20.03.2023)

Figure 9. Cluster Map (Created at 20.03.2023)
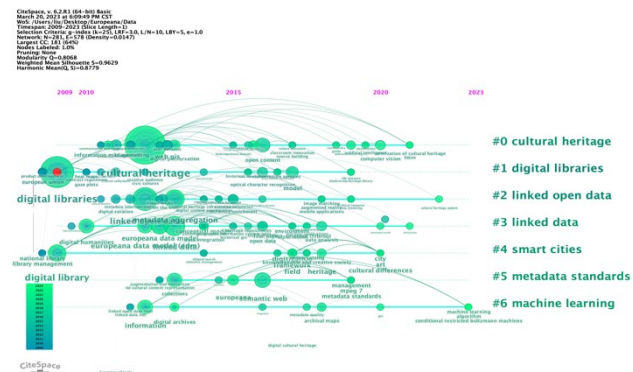

Figure 10. Timeline of Cluster (Created at 20.03.2023)

Fig 9 and Fig 10 illustrate keywords clustering and their timeline. Seven clustering results are shown, namely cultural heritage, digital libraries, linked open data, linked data, smart cities, metadata standards and machine learning. Digital libraries and smart cities were the first to appear, both stopping around 2020. Machine learning and Linked Open Data appeared in 2012 and are the longest-lasting clusters, still prevalent today. On the other hand, metadata standards were the latest to appear and earliest to end, with a very short research window.

3.1 Technical Framework Construction

The technical framework construction of Europeana is based on a distributed system architecture that allows for the aggregation and interoperability of digital cultural heritage content from multiple institutions across Europe. Europeana collects metadata from cultural heritage institutions through a process known as metadata contribution. The collections need to make a conscious

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

decision to make data available and program an export dataset compatible with the EDM. This process involves mapping the metadata to common data elements. The normalised metadata is then indexed and made available through Europeana's search and discovery interface, which allows users to find and retrieve digital objects based on specific criteria, such as keywords, date ranges, and media type.

At the beginning, the Europeana Semantic Elements (ESE) was used for Europeana, but it soon became clear that it would need to be replaced. From 2009, work began on the Europeana Semantic Data Layer, a task of the Europeana Connect project. As a result of the process of specifying an EDM (Bontchev, 2012), EDM was considered a candidate for operationalisation. EDM adheres to the Web of Data principles and is serialisable in a variety of XML and RDF syntaxes (i.e., N-Triples, Turtle, JSON-LD, etc.) (Freire et al., 2018). This flexibility allows and encourages technological innovation.

Europeana manages data through a centralised database system, which ensures the preservation and accessibility of the digitised cultural heritage content over time. The database is designed to support the efficient retrieval and delivery of digital objects to users. Europeana also promotes the use of Linked Open Data, which allows for the linking and reuse of cultural heritage data across the web. This enhances the discoverability and accessibility of cultural heritage content and supports the development of new applications and services based on the data.

3.2 Digital Resources

Europeana offers a wide variety of digital resources. It provides access to digitised books, manuscripts, and other written works from libraries and archives across Europe, including valuable insights into Europe's cultural heritage. Also, Europeana gives access to a wide range of images, including photographs, paintings, drawings, and other forms of visual material. These resources form a visual representation and a unique perspective on European history and art. Moreover, audio and video resources, including recordings of music, sound archives, and film as well as digitised museum objects, including artifacts, sculptures, and other objects of material culture, are also available through Europeana. When it launched in November 2008, Europeana gave users access to 4.5 million digital objects, and this total is increasing every year. As of early 2023, it provides access to over 55 million digital objects (*Brief History*, 2023). Europeana does not contain the actual digital objects, which are hosted by the owning institution, instead, it contains pointers to the objects in the form of metadata records (MR). MRs are used for most of the services that Europeana provides, such as searching and browsing (Berardi et al., 2012).

Europeana provides access to a wealth of digital cultural heritage resources from across Europe and is designed to support the preservation, dissemination, and use of this valuable cultural heritage for generations to come. Moreover, Europeana is intended to promote Europe's digital economy, for example, holding online exhibitions and participating in crowd-sourcing experiments (Nicholas & Clark, 2013).

3.3 Service System

Target Users: Europeana's target users are segmented into five categories: general users with an interest in culture or history, school students and academic users, both using the resources in (albeit widely differing) educational contexts, expert users interested in specific topics, and professional users, probably employed by CH and GLAM institutions (Purday, 2009b).

Search and Discovery: Users search and retrieve digital cultural heritage resources based on specific criteria, such as keywords, date ranges, and media type. The search results can be filtered and refined, making it easy for users to find the needed resources. Europeana's search architecture is based on Apache Solr, a widely-implemented open-source search platform (Petras et al., 2017). Most search requests (ca. 70%) are for named entities, in particular geographic searches and people or Cultural Heritage objects. Although English is the most popular language on the site, people also use their native tongue (Petras et al., 2017).

Europeana Website Homepage: In the middle of the web interface is a search bar. In addition, there is a "Collections" menu at the top right, which leads to seven options: Theme, Topics, Features, Centuries, Galleries, Organisations and Recent items. Below each collection are some specific examples, such as Colouring Books under Features (Fig 12). Moreover, in the "Stories" (Fig 11) option, there are specially curated selections from Europeana's collections. There are currently more than 900 items available (*Stories*, 2023). The "For professionals" is a link to the "Europeana Pro" website with descriptions.
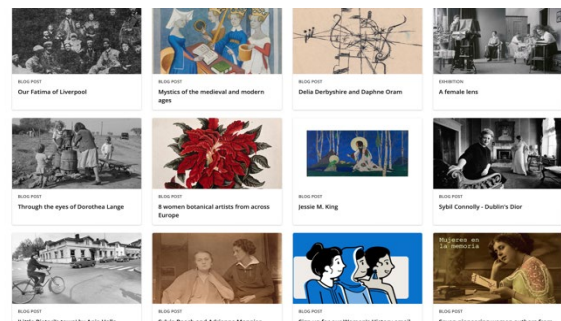

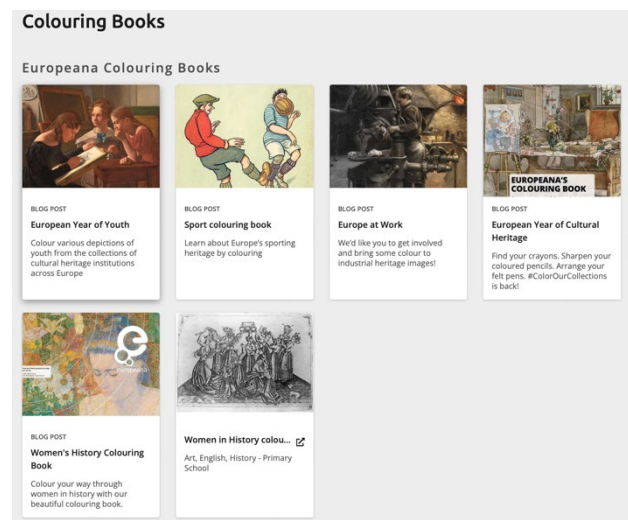**Figure 11**. Europeana Stories (*Stories*, 2023)

Figure 12. Colouring Books (Colouring Books, 2023)

Alongside basic searching and browsing, another aspect of Europeana's functionality is **Data Reuse**: the affordance of reusable cultural heritage through the provision of open and machine-readable data, which can be used by researchers, educators, and developers to build new applications and services. Europeana also provides tools and services that support the development of new applications, including APIs and software development kits (Charles & Isaac, 2015). In addition, Europeana supports education and outreach activities by providing resources and tools for both educators and students, including lesson plans, educational videos, and online games. Europeana also provides outreach materials, such as brochures and posters, to help promote the platform and its resources. Furthermore, Europeana

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

provides technical support to its users, including support for data ingestion and normalisation, as well as technical support for data retrieval and display. As mentioned above, Europeana also provides support for the development of novel applications and services which use its data. At the same time, evaluations have aided the development of Europeana, though specific implementations cannot be linked to specific evaluation findings (Petras & Stiller, 2017).

4. SAMPO MODEL

The Sampo Model is a concept developed by the Semantic Computing Research Group (SeCo) from Aalto University and University of Helsinki in Finland (*Semantic Computing Research Group (SeCo)*, 2023). "Sampo" refers to a magical artifact in Finnish mythology that brings wealth and good fortune.

Following are some basic analyses of Sampo from Citespace, with the search topic TS=(Sampo Model). The total publications (Fig 13) are much smaller than those for Europeana or CIDOC CRM, with a total of 16 articles up to 2022, with some years showing no publications at all. Because of the very limited amount of data, it is impossible to perform cluster analysis.
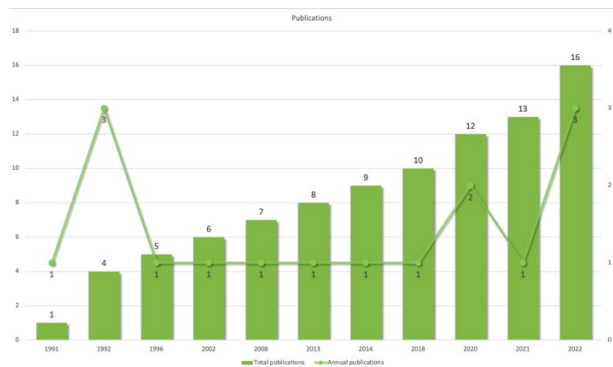

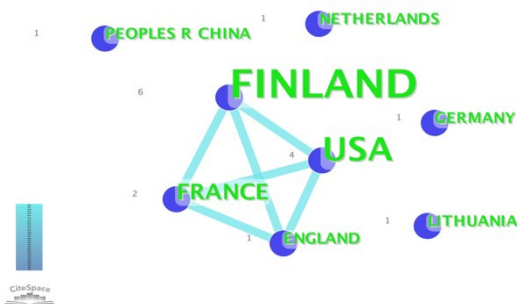**Figure 13.** Publications of Sampo Model
(Created at 20.03.2023)

Figure 14. Country of Publications (Created at 20.03.2023)

Fig 14 shows a Country Analysis Map of publications. The largest source of articles (six) is Finland, followed by the USA with four. The graph also shows collaborations between Finland, USA, France, and England, while the other countries shown have published articles independently.

The Sampo Model refers to the policies, regulations, and norms that shape the cultural environment in Finland. This includes

factors such as the Finnish education system, emphasising the importance of collaboration and networking within the Finnish cultural and academic world. Several portals have been developed based on the Sampo Model. The following analysis uses one such portal, WarSampo (*WarSampo*, 2023), as an example.

4.1 Technical Framework Construction

WarSampo is a digital humanities project that aims to provide a comprehensive understanding of the impact of World War II on Finland and Finnish society. It has been online since 2015 with several new application perspectives published in 2016–2019. The WarSampo application won the LODLAM Open Data Prize in Venice in 2017 (Hyvönen, 2022). The project uses various sources such as war diaries, official documents, photographs, maps, and other archival materials to create a rich, interactive database of the war years.

The WarSampo project was initiated by the University of Helsinki in 2017 and has since grown into a collaborative effort involving several Finnish institutions, including the Finnish National Archives, the Finnish Defence Forces, and the Finnish Broadcasting Company. WarSampo consists of two parts: First, a Linked Open Data service called WarSampo Data enables Digital Humanities study and application development on the history of the war. Second, a semantic WarSampo Portal has been developed to evaluate and demonstrate the data service's usefulness. WarSampo is one of the first large-scale systems for serving and publishing WW2 LOD on the Semantic Web (Hyvönen et al., 2016).

WarSampo is built on a combination of tools and technologies based on the principles of Linked Open Data, a common framework for publishing and connecting data on the web. The LOD approach enables data integration and interoperability between different datasets, which is essential for creating a comprehensive database of the war years. WarSampo uses ontologies to create a formalised and standardised representation of the domain knowledge. The project is based on a range of tools and techniques for extracting data from various sources, including digitised archives, official documents, and war diaries. The various visualisation tools help users explore and understand the data in the database, including maps and timelines.

4.2 Digital Resources

The WarSampo database includes information on various aspects of the war, such as military operations, civilian experiences, propaganda, and post-war reconstruction efforts, with a core dataset of the casualty register of the National Archives of Finland (Koho and Hyvönen, 2022). The project also incorporates advanced data visualization and text-mining techniques to help users better understand the complex socio-political dynamics of war. The Finnish Winter War (1939–1940) against the Soviets, the Continuation War (1941–1944), in which the Finns briefly retook the captured Winter War lands, and the Lapland War (1944–1955), in which the Finns drove the Germans from Lapland, are all covered by the WarSampo Data Service. The datasets that were used are shown in Fig 15 (Hyvönen et al., 2016).

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

#	Name	Providing organization	Size
1	Casualties of WW2	National Archives	94,700 death records
2	War diaries	National Archives	13,000 war diaries of troops
3	Photos & films	Defence Forces	160,000 photos & films
4	Kansa Taisteli magazine articles	The Assoc. for Military History in Finland & Bonnier	3,357 articles of veteran soldiers
5	Karelian places	Jyrki Tiittanen / National Land Survey	32,400 places of the annexed Karelia
6	Karelian maps	National Land Survey	47 wartime maps of Karelia
7	Senate atlas	National Archives	404 historical maps of Finland
8	Municipalities	National Archives	625 wartime municipalities
9	Organization cards	National Archives	ca 500 army units & ca 300 persons & 642 battles
10	National Biography	Finnish Literature Society	ca 500 biographies of wartime persons
11	Wartime events	War history books	1,000 events
12	Persons	War history books, Wikipedia	2,600 persons
13	Army units	War history books	3,200 army units

Figure 15. Datasets of WarSampo (Hyvönen et al., 2016)
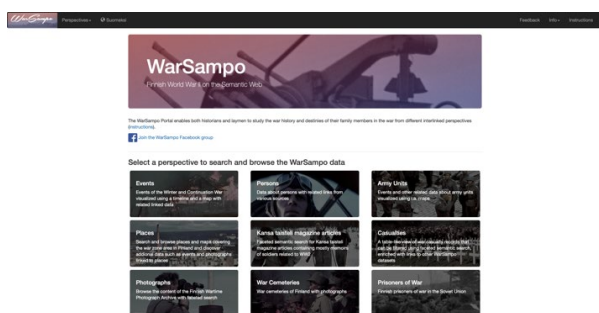

Figure 16. Homepage of WarSampo

Fig 16 shows the homepage of WarSampo and nine interlinked application perspectives for accessing the underlying data: Major events (1050), People (100000), Army Units (15900), Places (an alternative view for events), Kansa Taisteli magazine articles (3360), Casualty data (95000 death records), Authentic photographs (160000), War Cemeteries (630) and Finnish Prisoners of War (4500) (Hyvönen, 2022). The data is presented in a multitude of formats. For example, events are depicted on a timeline, on maps, and with connected data; people are linked with biographies and further information. Events, war diaries, and persons connected to Army formations are also included, and both contemporary and historic places provide a geographic perspective for looking up events in the battle zones. Articles from the Kansa Taisteli magazine include soldiers' recollections and there are photographs from the front lines and the War Cemetery of the World War II Dead (Hyvönen, 2022).

4.3 Service System

The Sampo Model is based on six principles: 1. Support collaborative data creation and publishing. 2. Use a shared open ontology infrastructure. 3. Support data analysis and knowledge discovery in addition to data exploration. 4. Provide multiple perspectives to the same data. 5. Standardise portal usage by a simple filter-analyse two-step cycle. 6. Make clear distinction between the Linked Open Data service and the user interface.

The Linked Data Finland platform provides Sampo portal data services and since 2019 the Sampo-UI framework has been used for user interfaces (Hyvönen, 2022). Based on this, the SeCo team has already developed over 30 sub-projects. The Sampo Model is extensible and can be used for visualisation of any field with a dataset. One of the innovations of the Sampo model and portals is to offer the user tools to analyse and visualise more

detailed information than is possible with basic searching and browsing (Hyvönen, 2020). Given the scale and complexity of the knowledge graph, it would be expensive and time-consuming to edit the data for each individual set of user requirements; instead, WarSampo allows multiple application viewpoints to be supported merely by changing how the data is accessible via SPARQL. This makes adding new application viewpoints to the data simpler and more independent without impacting the existing perspectives (Hyvönen et al., 2016).

5. CONCLUSION

In the digital age, different users have widely differing needs, and in the cultural heritage field, different LOD models have been developed to meet these needs. This review focusses on CIDOC CRM, Europeana and Sampo Model as examples, though future work will include further examples such as World Digital Library, HathiTrust, Linked Art Model, CrossCult project etc., as well as applications and systems such as the LOD-based vocabulary at Bamberg University (*Vokabular Dienst: BVHA: Alphabetical Index:*, 2023).

The analysis of publications, shows that, although Europeana appeared more recently than CIDOC CRM, it has been studied more frequently, though research on both peaked around the same time in the early 2010s. (In contrast, the Sampo model is more localised, with the majority of its 16 publications coming from Finland). Overall, this indicates that digital model research reached its peak almost a decade ago.

While it should be noted that not everyone is positive about digital models, with some researchers and conservators holding the opinion that the effort required to set them up and operate them outweighs their benefits (Arera-Rütenik, 2021), we believe that each of the models examined in this paper has its own unique strengths. CIDOC CRM is a model that provides a common and formal language for describing cultural heritage objects and their relationships. As a framework for organising and structuring cultural heritage data, its ability to represent complex relationships and hierarchies among cultural heritage objects in particularly strong.

Europeana is a digital platform that provides access to millions of digitised cultural heritage resources from European GLAM and cultural institutions, giving users a single access point to explore and discover Europe's cultural heritage. Its strength lies in its vast and diverse collection of objects, which span multiple disciplines and domains. Europeana also provides powerful search and discovery tools, enabling users to easily explore and discover cultural heritage resources.

The Sampo Model is a digital infrastructure for managing and presenting cultural heritage data in Finland. It consists of modular tools and services that can be customised and combined to create tailored digital heritage applications. The Sampo Model is powerful in its flexibility and modularity, allowing the creation of bespoke digital heritage applications that meet specific needs in multiple domains, including eCulture, eHealth, eLearning, eGovernment, eCommerce, eGeography, and eBiology. Currently the Sampo Model is only in use in Finland but the concepts could easily be adapted for use elsewhere.

REFERENCES

Arera-Rütenik, T. (2018). Digital Humanities in der Bauforschung: Systematik und Potenzial kodierter Bau- und Befundbeschreibungen als Wissensgewinnungs und Wissensdistributionswerkzeug. In S. Breitling & J. Giese

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

(Eds.), Bauforschung in der Denkmalpflege: Qualitätsstandards und Wissensdistribution. University of Bamberg Press.

Berardi, G., Esuli, A., Gordea, S., Marcheggiani, D., & Sebastiani, F. (2012). Metadata Enrichment Services for the Europeana Digital Library. In P. Zaphiris, G. Buchanan, E. Rasmussen, & F. Loizides (Eds.), *Theory and Practice of Digital Libraries* (Vol. 7489, pp. 508–511). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33290-6_61

Bontchev, B. (2012). Evolving Europeana 's Metadata: From ESE to EDM. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 2, 27–37. https://doi.org/10.55630/dipp.2012.2.12

Brief History. (2023). Europeana Pro. https://pro.europeana.eu/about-us/mission

Čerāns, K., Ovčiņņikova, J., Bojārs, U., Grasmanis, M., Lāce, L., & Romāne, A. (2021). Schema-Backed Visual Queries over Europeana and Other Linked Data Resources. In R. Verborgh, A. Dimou, A. Hogan, C. d'Amato, I. Tiddi, A. Bröring, S. Mayer, F. Ongenae, R. Tommasini, & M. Alam (Eds.), *The Semantic Web: ESWC 2021 Satellite Events* (Vol. 12739, pp. 82–87). Springer International Publishing. https://doi.org/10.1007/978-3-030-80418-3_15

Charles, V., & Isaac, A. (2015). *Enhancing the Europeana Data Model (EDM)*.

Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences*, 101(suppl_1), 5303–5310. https://doi.org/10.1073/pnas.0307513100

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377. https://doi.org/10.1002/asi.20317

Classes & Properties Declarations of CIDOC-CRM version: 7.1.2. (2023). https://www.cidoc-crm.org/html/cidoc_crm_v7.1.2.html

Colouring Books. (2023). https://www.europeana.eu/en/colouring-books

Freire, N., Charles, V., & Isaac, A. (2018). Evaluation of Schema.org for Aggregation of Cultural Heritage Metadata. In A. Gangemi, R. Navigli, M.-E. Vidal, P. Hitzler, R. Troncy, L. Hollink, A. Tordai, & M. Alam (Eds.), *The Semantic Web* (Vol. 10843, pp. 225–239). Springer International Publishing. https://doi.org/10.1007/978-3-319-93417-4_15

Haller, A., & Polleres, A. (2020). Are we better off with just one ontology on the Web? *Semantic Web*, 11(1), 87–99. https://doi.org/10.3233/SW-190379

Hiebel, G., Doerr, M., & Eide, Ø. (2017). CRMgeo: A spatiotemporal extension of CIDOC-CRM. *International Journal on Digital Libraries*, 18(4), 271–279. https://doi.org/10.1007/s00799-016-0192-4

Hyvönen, E. (2020). Using the Semantic Web in digital humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery. *Semantic Web*, 11(1), 187–193. https://doi.org/10.3233/SW-190386

Hyvönen, E. (2022). Digital humanities on the Semantic Web: Sampo model and portal series. *Semantic Web*, 1–16. https://doi.org/10.3233/SW-223034

Hyvönen, E., Heino, E., Leskinen, P., Ikkala, E., Koho, M., Tamper, M., Tuominen, J., & Mäkelä, E. (2016). WarSampo Data Service and Semantic Portal for Publishing Linked Open Data About the Second World War History. In H. Sack, E. Blomqvist, M. d'Aquin, C. Ghidini, S. P. Ponzetto, & C. Lange (Eds.), *The Semantic Web. Latest Advances and New Domains* (Vol. 9678, pp. 758–773). Springer International Publishing. https://doi.org/10.1007/978-3-319-34129-3_46

Koho, M., & Hyvönen, E. (2022). *Studying Occupations and Social Measures of Perished Soldiers in WarSampo Linked Open Data*.

Mazurek, C., Sielski, K., Walkowska, J., & Werla, M. (2012). *Applicability of CIDOC CRM in Digital Libraries*.

Niccolucci, F. (2017). Extending, mapping, and focusing the CIDOC CRM: Introduction to the special issue. *International Journal on Digital Libraries*, 18(4), 251–252. https://doi.org/10.1007/s00799-016-0198-y

Nicholas, D., & Clark, D. (2013). The second digital transition: To the mobile space – an analysis of Europeana. *Learned Publishing*, 26(4), 240–252. https://doi.org/10.1087/20130402

Oldman, D. (2014). *The CIDOC Conceptual Reference Model (CIDOC-CRM): PRIMER*.

Petras, V., Hill, T., Stiller, J., & Gäde, M. (2017). Europeana – a Search Engine for Digitised Cultural Heritage Material. *Datenbank-Spektrum*, 17(1), 41–46. https://doi.org/10.1007/s13222-016-0238-1

Petras, V., & Stiller, J. (2017). A Decade of Evaluating Europeana—Constructs, Contexts, Methods & Criteria. In J. Kamps, G. Tsakonas, Y. Manolopoulos, L. Iliadis, & I. Karydis (Eds.), *Research and Advanced Technology for Digital Libraries* (Vol. 10450, pp. 233–245). Springer International Publishing. https://doi.org/10.1007/978-3-319-67008-9_19

Purday, J. (2009a). Think culture: Europeana.eu from concept to construction. *BIBLIOTHEK Forschung Und Praxis*, 33(2). https://doi.org/10.1515/bfup.2009.018

Purday, J. (2009b). Think culture: Europeana.eu from concept to construction. *The Electronic Library*, 27(6), 919–937. https://doi.org/10.1108/02640470911004039
Semantic Computing Research Group (SeCo). (2023). https://seco.cs.aalto.fi/

Stories. (2023). https://www.europeana.eu/en/stories
Tzitzikas, Y., Mountantonakis, M., Fafalios, P., & Marketakis, Y. (2022). CIDOC-CRM and Machine Learning: A Survey and Future Research. *Heritage*, 5(3), 1612–1636. https://doi.org/10.3390/heritage5030084

Use Cases | CIDOC CRM. (2023). https://www.cidoc-crm.org/useCasesPage

Vokabular Dienst: BVHA: Alphabetical index: (2023). https://hist-arch-vocab.org/skosmos/bvha/en/index

WarSampo. (2023). https://www.sotasampo.fi/en/